Probability theory of capacity value of additional generation

Citation for published version:
Zachary, S & Dent, C 2012, 'Probability theory of capacity value of additional generation', *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 226, no. 1, pp. 33-43. https://doi.org/10.1177/1748006X11418288

Digital Object Identifier (DOI):
10.1177/1748006X11418288

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability

Probability theory of capacity value of additional generation

S. Zachary* and C.J. Dent†

May 22, 2011

Abstract

This paper describes a new probability theory of the capacity value of additional generation in electrical power systems. A closed form expression for the ELCC (Effective Load Carrying Capability) or EFC (Equivalent Firm Capacity) of a small additional capacity is derived. This depends on the mean and variance of the distribution of available additional generation capacity, and the shape of the distribution of the difference between available existing capacity and demand, near zero margin. The theory extends naturally to the case where the pre-existing background and additional resource are not statistically independent.

The theory may be used to explain and confirm the generality of various well-known properties of capacity value results, as is illustrated

*Department of Actuarial Mathematics and Statistics, Heriot-Watt University, U.K.

†Corresponding author. Address: School of Engineering and Computing Sciences, Durham University, South Road, Durham DH1 3TX, U.K. Email: chris.dent@durham.ac.uk.

using Great Britain examples. Of particular note is the common observation that if the distribution for demand is shifted so as to increase the calculated risk, then the capacity value of additional generation increases. The new theory demonstrates that this is not true in general, but rather is a consequence of the shape near zero margin of the probability distribution of the margin of existing generating capacity over demand.

1 Introduction

Electrical energy supply should at all times be sufficient to meet demand, as opportunities for large-scale storage are very limited with present technologies. However, since both supplies and demands are uncertain, there is always a small *Loss of Load Probability* (LOLP) that supply will be insufficient to meet demand.

Renewable sources of supply, of which installed capacities are increasing rapidly in power systems worldwide, are naturally uncertain as their availability to generate depends on the availability of the relevant natural resource[1]. This is in contrast to conventional generation (coal, gas, nuclear etc.), whose availability to generate given adequate fuel supplies is primarily a matter of mechanical availability.

Given this qualitative difference, it is helpful to define an effective capacity (usually called *capacity value* or *capacity credit*, the terms are used interchangeably) which measures in some sense the contribution of the additional resource. Such capacity values have been used widely in practical wind integration studies, for instance [1, 2]; an IEEE Power and Energy So-

[1] While tides are predictable many years ahead, the coincidence of a large tidal resource with periods of very high demand remains uncertain far ahead of real time.

ciety task force recently published a comprehensive review [3] of calculation approaches and applications.

Various definitions of capacity value are possible. A comparison of results calculated using different definitions may be found in [4]. This paper will consider the two most common ones, namely *Effective Load Carrying Capability* (ELCC, the additional demand which the additional generation can support without increasing risk), and *Equivalent Firm Capacity* (EFC, the deterministic capacity whose addition would give the same decrease in risk). The generation adequacy risk is typically measured by *Loss of Load Probability* (LOLP, the probability that supply is insufficient to meet demand at a particular time), or *Loss of Load Expectation* (LOLE, the expected number of periods during a given time window in which supply is insufficient to meet demand.)

This paper will show that ELCC and EFC are mathematically the same, apart from a shift of reference point and a sign change—an observation which will be important for the subsequent probability theory. It should also be noted that the concept of capacity value is only well-defined in relation to the stochastic background created by the pre-existing supplies and demands.

A closed form expression (essentially the leading terms of a Taylor series) for the capacity value will be derived in the case where the uncertainty in the added capacity (as measured for example by its standard deviation) is small in relation to that already to be found in the existing supply and demand; ELCC and EFC are then, for practical purposes, coincident. The special case of a normal distribution for the distribution of the margin of existing supply over demand is known in the power systems literature as the *z-method* [5]; however, differently from previous analytical expressions for capacity values, these new results make no assumptions about the form of the various distributions, the derivation presented will be explicit about the limited assumptions required, and the theory extends naturally to the case where the pre-existing background and additional resource are not statistically independent.

As expected the zeroth-order term in the capacity value expression is just the mean of the distribution of the additional capacity, while the first-order term represents a correction proportional to its variance. Higher order terms (negligible when the variation in the available capacity from the new generation is small in relation to the variation in the margin of existing supply over demand) are in principle calculable, fairly readily when the additional capacity is statistically independent of the existing capacity and demand. However, their messy algebraic derivation provides no further insight, and (where effects of finite network capacity are not considered) it is straightforward to obtain numerical results.

Section 2 will describe the probability theory of capacity values (its application to this area being new). Section 3 will then show how this theory maps into a range of practical capacity value calculations involving different underlying risk measures, and the insights which result. Finally conclusions will be presented in Section 4.

2 Probability theory of capacity values

2.1 Definitions

Consider initially the state of a system at a given time.

Let the random variable M denote the excess of supply (for example, conventional generation plus existing wind generation) over demand in an existing system, and suppose that M has a (cumulative) distribution function F_M with associated, and differentiable, density f_M. Then the associated *risk* is the *loss of load probability* (LOLP):

$$\mathbf{P}(M \leq 0) = F_M(0).$$

Suppose now that a further capacity is added, represented by the random variable Y (which may or may not be independent of M) with distribution function F_Y. Suppose also that $\mathbf{E}Y = \mu_Y$ and $\mathbf{var}Y = \sigma_Y^2$.

The two most commonly used definitions of the *capacity value* of Y are:

Effective load carrying capacity (ELCC): This is given by the solution ν_Y^{ELCC} of

$$\mathbf{P}(M + Y - \nu_Y^{\text{ELCC}} \leq 0) = \mathbf{P}(M \leq 0) = F_M(0), \tag{1}$$

i.e. the amount of further demand which may be added while maintaining the same level of risk.

Equivalent firm capacity (EFC): This is given by the solution ν_Y^{EFC} of

$$\mathbf{P}(M + Y \leq 0) = \mathbf{P}(M + \nu_Y^{\text{EFC}} \leq 0) = F_M(-\nu_Y^{\text{EFC}}), \tag{2}$$

i.e. the amount of deterministic capacity ν_Y^{EFC} whose addition would result in the same level of risk as that of the random capacity Y.

Note that in the case where Y is deterministic ($Y = \mu_Y$ always) we have $\nu_Y^{\text{ELCC}} = \nu_Y^{\text{EFC}} = \mu_Y$. More generally both ν_Y^{ELCC} and ν_Y^{EFC} depend on the distributions of both M and Y, and it is one purpose of this paper to explore this relationship.

Remark 1. Observe also that it follows from (1) and (2) that the EFC ν_Y^{EFC} is also the ELCC in the case where the surplus random variable M is replaced by $M + \nu_Y^{\text{EFC}}$. In particular we shall subsequently make use of this observation to translate immediately results obtained for ELCCs to their equivalents for EFCs.

2.2 Capacity values for small additions

Suppose now that the *variation* in Y, as measured for example by its standard deviation σ_Y is *relatively* small in comparison to the *variation* in M. (This will frequently be the case in applications as both existing capacity and demand are subject to considerable uncertainty.) It is shown here that both the ELCC and EFC are given by (typically very) small corrections to μ_Y each such correction being a multiple of σ_Y^2. (To simplify understanding of the arguments below, and to check that they may be made rigorous with regard to the dropping of small-order terms, e.g. in Taylor expansions, it may be supposed that units are chosen so that M is of unit size, while the variation in Y is of order $\epsilon \ll 1$, and hence σ_Y^2 is of order ϵ^2. In the application of results there is of course no need to actually choose such units.) It is further shown that when additionally μ_Y itself is relatively small, the ELCC and EFC essentially coincide.

The first case below assumes that M and Y are independent; the adjustments required for the general case are then derived.

2.2.1 The case where M and Y are independent

The following result gives the capacity value for this case.

Result 1. *Assume M and Y are independent, and the distribution function $F_M(m)$ is left-tail exponential for all $m \leq 0$ (i.e. $F_M(m) = ce^{k_M m}$). Then the distribution of $M' = M + W$ is $F_{M'}(m') = F_M(m' - \nu)$, where*

$$\nu = -\frac{1}{\lambda_M} \ln \left[\mathrm{d}w \, e^{-k_M w} f_W(w) \right]. \tag{3}$$

Proof. Rewrite (1) as

$$\int_{\mathbb{R}} \mathrm{d}F_Y(y) F_M(\nu_Y^{\mathrm{ELCC}} - y) = F_M(0). \tag{4}$$

(In the case where the distribution of Y has a density f_Y the term $\mathrm{d}F_Y(y)$ in the above integral may be more familiarly written as $f_Y(y)\mathrm{d}y$.) From the Taylor expansion of F_M about the critical point 0 it follows that

$$F_M(y) = F_M(0) + f_M(0)y + \frac{f'_M(0)}{2}y^2 + o(y^2), \tag{5}$$

as $y \to 0$. Now recall that when $\sigma_Y = 0$ (Y is deterministic), $\nu_Y^{\text{ELCC}} = \mu_Y$. More generally, under our present assumption that σ_Y is small, $\nu_Y^{\text{ELCC}} - \mu_Y$ is small, and so, in (4), $y - \nu_Y^{\text{ELCC}}$ is small (of order ϵ in the units of the discussion above). The $o(y^2)$ term in (5) may then be ignored, and the following is obtained by substitution into (4):

$$\int_{\mathbb{R}} \mathrm{d}F_Y(y) \left(F_M(0) + f_M(0)(\nu_Y^{\text{ELCC}} - y) + \frac{f'_M(0)}{2}(\nu_Y^{\text{ELCC}} - y)^2 \right) = F_M(0).$$

Simplifying,

$$\begin{aligned} f_M(0)(\mu_Y - \nu_Y^{\text{ELCC}}) &= \frac{f'_M(0)}{2}\mathbf{E}\left((Y - \nu_Y^{\text{ELCC}})^2\right) \\ &= \frac{f'_M(0)}{2}(\sigma_Y^2 + (\mu_Y - \nu_Y^{\text{ELCC}})^2). \end{aligned} \tag{6}$$

Recalling again that σ_Y^2 is small in relation to $f_M(0)/f'_M(0)$, it is straightforward that the quadratic equation (6) in $\mu_Y - \nu_Y^{\text{ELCC}}$ has the solution

$$\nu_Y^{\text{ELCC}} = \mu_Y - \frac{f'_M(0)}{2f_M(0)}\sigma_Y^2 + O(\sigma_Y^4),$$

as $\sigma_Y^2 \to 0$ so that (3) follows as required. □

Remark 2. (In the units suggested at the start of this section) the second "correction" term on the right side of (3) is of order ϵ^2, and the error in (3) as the solution of the quadratic equation (6) is of order ϵ^4. However, a more careful analysis of the Taylor expansion used in the above proof shows that, at least under very mild regularity conditions, the overall error in (3) is of order ϵ^3.

Remark 3. It follows immediately from the observation of Remark 1 (or by a repetition of the above analysis), that, under the same conditions as those of Result 1, the corresponding approximation for the EFC ν^{EFC} is given by

$$\nu_Y^{\mathrm{EFC}} = \mu_Y - \frac{f'_M(-\nu_Y^{\mathrm{EFC}})}{2f_M(-\nu_Y^{\mathrm{EFC}})}\sigma_Y^2. \tag{7}$$

In particular when the mean μ_Y of Y, and so also ν_Y^{EFC}, is small the EFC and the ELCC essentially coincide.

2.2.2 The case where M and Y may be dependent

Recall that the variation of Y about its mean value μ_Y is assumed to be small in relation to the variation in M. Thus the major contribution to the probability given by the left side of (1) is from values of M in the neighbourhood of 0. In the case where M and Y are not necessarily independent it is natural to simply modify the expression (4) to use the conditional distribution of Y given $M = 0$. Then Result 1 remains as before except only that μ_Y and σ_Y^2 are replaced by $\mu_{Y|M=0}$ and $\sigma^2_{Y|M=0}$, the respective mean and variance of the distribution of Y *conditional on* $M = 0$. (It is easy to check, for example, that this represents the required modification in the multivariate normal case—see below.)

Thus the distribution of the additional capacity corresponding to those times when the system is critically loaded should be used, as would be expected.

2.2.3 Further observations

The following further observations, for the case where M and Y are independent, all follow from Result 1. The corresponding observations for the "non-independent" case, again whenever the variation in Y is small in relation to that in M, are obtained as usual by using instead the conditional distribution of Y given $M = 0$.

1. The capacity value of the addition Y is its mean μ_Y, less a correction to allow for the additional uncertainty σ_Y (note that $f'_M(0)/f_M(0)$ will typically be positive). However, when the variation of Y about its mean is small (of order ϵ), then σ_Y^2—and so the correction given by Result 1—is actually very small (of order ϵ^2).

2. The correction from the mean identified above becomes more significant as σ_Y increases. In particular if μ_Y and σ_Y are scaled up in proportion to each other, as when the size of the additional resource is simply scaled up, then the *relative* correction

$$\frac{f'_M(0)}{2f_M(0)}\frac{\sigma_Y^2}{\mu_Y}$$

given by Result 1 grows in proportion to the scale factor.

3. If several *independent* small capacity increments Y_i, also independent of M, are made, then the capacity value of the total is the sum of the individual capacity values. (There is no contradiction here with the previous observation, since, as independent increments are added, the standard deviation of their sum, relative to the mean of the sum, decreases.) Thus, under the addition of independent resources, the usual benefits of statistical aggregation apply.

4. It is readily verified that, in general, when the left tail of the distribution of the surplus M is light-tailed, then

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{f'_M(x)}{f_M(x)} < 0 \tag{8}$$

(in the region of the left tail—indeed this inequality might almost be taken as a definition of "local" light-tailedness). It thus follows from Result 1 that if the distribution of M is *shifted* so as to decrease the risk, then the corresponding capacity value of an added resource decreases. This corresponds to often-made empirical observations, and to the result of direct calculation for the normal distribution. For a heavy-tailed distribution (very rare in practice in the present context) the inequality in (8) is in general reversed, and here if the distribution of M were shifted so as to decrease the risk, then the corresponding capacity value of an added resource would increase.

5. Finally, note that the shape of the distribution of Y (other than the assumption that its variance is relatively small) is irrelevant to Result 1 and its consequences.

2.2.4 Special cases for the distribution of M

1. *The Gaussian case.* Suppose that $M \sim N(\mu_M, \sigma_M^2)$, where $\mu_M > 0$. Then $f'_M(0)/f_M(0) = \mu_M/\sigma_M^2$, so that (3) becomes

$$\nu_Y^{\mathrm{ELCC}} = \mu_Y - \frac{\mu_M}{2\sigma_M^2}\sigma_Y^2. \tag{9}$$

This is precisely the result obtained from what is commonly known as the *z-method*, first derived in [5] under an explicit assumption that the distribution of available surplus capacity M does not change shape when further capacity Y is added. However, for this to be so essentially requires that the distribution of both M and Y be normal. This indeed is the only case under which we have the precise result given by (9) (as may be verified from the differential equation given by a comparison of (3) and (9)). Result 1 of the present paper gives the corresponding result for more general distributions of M and Y, subject only to the assumption made there as to their relative variation.

2. *The exponential-tail case.* Suppose that, in the region of $x = 0$, the density of M satisfies $f_M(x) \approx k \exp(\lambda x)$, for some constant k, as might approximately be the case if the *right* tail of the distribution of demand D were approximately exponential. Then (3) becomes

$$\nu_Y^{\text{ELCC}} = \mu_Y - \lambda\sigma_Y^2. \tag{10}$$

2.3 Further dependence results

Consider again the case where M and Y may be dependent. Here the use the modified version of Result 1 requires the conditional distribution of Y given $M = 0$; specifically it requires $\mu_{Y\,|\,M=0}$ and $\sigma^2_{Y\,|\,M=0}$ as defined earlier.

Suppose now that $M = X - D$ where, typically, X is existing generation and D is demand. Then the above calculations require knowledge of the joint distribution of X, D and Y. The calculations themselves are routine, if messy. However, there are two cases in which matters become slightly simpler.

The first case is where X is independent of the joint distribution of D and Y, as will frequently be the case if X represents conventional generation. The conditional distribution of Y given $M = 0$ is then given (in cumulative form) by

$$\begin{aligned}\mathbf{P}(Y \le y \,|\, M = 0) &= \int_{\mathbb{R}} \mathbf{P}(Y \le y \,|\, X = x, D = x)\, \mathrm{d}F_{X\,|\,M=0}(x)\\ &= \int_{\mathbb{R}} \mathbf{P}(Y \le y \,|\, D = x)\, \mathrm{d}F_{X\,|\,M=0}(x)\end{aligned} \tag{11}$$

where $F_{X\,|\,M=0}$ is the (cumulative) distribution function of X conditional on $M = 0$, and where the second line in the above expression use the assumed independence of X. (Where X has a corresponding conditional density $f_{X\,|\,M=0}$

$\mathrm{d}F_{X|M=0}(x)$ may be replaced in (11) by $f_{X|M=0}(x)\mathrm{d}x$.) The conditional mean $\mu_{Y|M=0}$ and variance $\sigma^2_{Y|M=0}$ of Y given $M = 0$ may then be calculated as usual. In particular, the following standard results may be derived from (11),

$$\begin{aligned}
\mu_{Y|M=0} &= \int_{\mathbb{R}} y \,\mathrm{d}\mathbf{P}(Y \le y \mid M = 0) \\
&= \int_{\mathbb{R}} y \left[\int_{\mathbb{R}} \mathrm{d}\mathbf{P}(Y \le y \mid D = x) \,\mathbf{d}F_{X|M=0}(x) \right] \\
&= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} y \,\mathrm{d}\mathbf{P}(Y \le y \mid D = x) \right] \mathbf{d}F_{X|M=0}(x) \\
&= \int_{R} \mu_{Y|D=x} \,\mathrm{d}F_{X|M=0}(x) \qquad (12)
\end{aligned}$$

and, similarly,

$$\mathbf{E}(Y^2 \mid M = 0) = \int_{R} \mathbf{E}(Y^2 \mid D = x) \,\mathrm{d}F_{X|M=0}(x)$$

so that

$$\sigma^2_{Y|M=0} = \mathbf{E}(Y^2 \mid M = 0) - \mu^2_{Y|M=0}. \qquad (13)$$

These expressions are natural when the joint distribution of Y and D is known (as well as the distribution of the independent random variable X). For example, in the case where X and D have respective densities f_X and f_D, the above expression (12) may be written as

$$\mu_{Y|M=0} = \frac{\displaystyle\int_{\mathbb{R}} \mu_{Y|D=x} \, f_X(x) f_D(x) \,\mathrm{d}x}{\displaystyle\int_{\mathbb{R}} f_X(x) f_D(x) \,\mathrm{d}x}.$$

The second case is where D is independent of the joint distribution of X and Y, as may again well be the case. Here the results are entirely analogous to those given above, with the roles of X and D simply interchanging.

2.4 Capacity values over time

Thus far the probability theory has been associated with a single "snapshot" in time. However, capacity values are typically defined to correspond to a extended period such as a year, where the relevant probability distributions vary from day to day, or hour to hour, within that extended period.

The appropriate measure of risk is now the sum of the LOLPs associated with the individual time periods within the longer period, i.e. the *loss of load expectation* (LOLE):

$$\sum_t \mathbf{P}(M_t \leq 0),$$

(where t indexes the individual time periods), and, for example, the definition (1) of the ELCC now becomes the solution ν_Y^{ELCC} of

$$\sum_t \mathbf{P}(M_t + Y_t - \nu_Y^{\mathrm{ELCC}} \leq 0) = \sum_t \mathbf{P}(M_t \leq 0) = \sum_t F_{M_t}(0). \tag{14}$$

Sometimes it may be sufficient to concentrate attention on particular "risky" hours or days over which the relevant random variables (defining M_t and Y_t) may be considered to be identically distributed, in which case the "snapshot" capacity value becomes the "extended" capacity value.

Otherwise one must work directly with, for example, equation (14). Under the analogous conditions to those of Result 1, i.e. that, for all t, the variance $\sigma^2_{Y_t}$ is small in relation to that of M_t and that M_t and Y_t are independent, then, to a similarly good approximation,

$$\nu_Y^{\mathrm{ELCC}} = \frac{1}{\sum_t f_{M_t}(0)} \left(\sum_t f_{M_t}(0)\mu_{Y_t} - \frac{1}{2} \sum_t f'_{M_t}(0)\sigma^2_{Y_t} \right). \tag{15}$$

The proof is as that of Result 1 with obvious modifications, and the result does of course reduce to that of Result 1 in the identically distributed case. Note also that the expression (15) simplifies considerably in the case where the Y_t are identically distributed.

In the case where M_t and Y_t fail to be independent, one again may work with conditional distributions.

2.5 Estimation of capacity values

The determination of capacity values requires knowledge of the associated probability distributions, and thus requires the availability of appropriate data.

In the "snapshot" case where the ELCC is to be estimated for a given time, sufficient knowledge is required of the joint distribution of the surplus M and the added capacity random variable Y. Ultimately this will have to be obtained from observation: for example repeated values of the triple (X_t, D_t, Y_t) might be observed over a sufficient number of time periods for which this triple may be considered identically distributed, the best inference being obtained in the case where there is also independence over time periods.

In the more general case where there is a requirement to estimate the ELCC associated with an extended period of (for example) a future year or peak season, then if relevant historical data are available a similar procedure may be followed. Suppose again that, for each future t, the random variable $M_t = X_t - D_t$ where X_t represents existing generation capacity and D_t represents demand. Rewrite (14) as

$$\sum_t \mathbf{P}(X_t - D_t + Y_t - \nu_Y^{\mathrm{ELCC}} \leq 0) = \sum_t \mathbf{P}(X_t - D_t \leq 0). \qquad (16)$$

Suppose now that the random variables X_t may be considered independent of the random variables D_t and Y_t, and that the distribution of the former is known (for example, from a knowledge of the probabilities of outages of individual generation plant). Suppose also that observed historic values (d_t, y_t) of the pair (D_t, Y_t) are available from one or more years (recall that typically Y_t is some form of renewable generation such as wind). Then the associated ELCC ν_Y^{ELCC} may be estimated as the solution of

$$\sum_t \mathbf{P}(X_t - d_t + y_t - \nu_Y^{\mathrm{ELCC}} \leq 0) = \sum_t \mathbf{P}(X_t - d_t \leq 0), \tag{17}$$

where the probability measure $\mathbf{P}$ now simply models the distributions of the random variables X_t (frequently these random variables will be considered independent and identically distributed), the sums are over historic times, and the (d_t, y_t) are rescaled according to projected future peak demand and additional installed generating capacity. This is the so-called *hindcast* approach, described in an application context in [3]. Its validity, in particular the *consistency* of the estimated ELCC, may be justified by an appeal to a fairly general version of the Central Limit Theorem.

3 Application to practical capacity value calculations

3.1 Great Britain data

The insights available from the above theory will be illustrated using sample data from Great Britain (GB). The discussion will be set in the context of well-known properties of capacity value calculations, which were surveyed in [4] and which have not previously been explained fully in terms of the underlying probabilistic model, as well as observations arising from the new theory. In all cases, the capacity value of a wind generation fleet, when added to an existing background of all-conventional generating plant, will be considered. The probability distribution of available conventional capacity will thus be assumed to be independent of both demand and the available wind capacity.

Coincident hourly time series for transmission-metered wind *load factor* [6] (LF, the wind output as a proportion of maximum) and demand [7] are available for Great Britain for the four winters between 2006 and 2010. The wind time series may be rescaled to a projection of wind power output for any given scenario of installed capacity, under an assumption that the probability distribution of load factors remains unchanged over time. An equivalent scaling may be made for demand using each winter's *Average Cold Spell* (ACS) peak demand[2]. There are some difficulties of comparability between years in both these rescalings of historic metered data, as both the geographical distribution of wind farms and underlying demand patterns change over time; in particular, an increasing amount of distribution-connected wind generation appears as negative demand in GB transmission-metered data from recent years. For practical wind integration studies, greater care in the use of historic data must therefore be taken, however for this illustrative purpose the simple approach taken is quite satisfactory.

The probability distribution of available conventional capacity is based on the list of units connected to the GB system in winter 2008/09[3]. Each individual unit is modelled as a 2-state random variable, with either 0 or max-

[2]ACS peak demand is the standard measure of underlying peak demand level in Great Britain, independent of the weather conditions in the year in question. For a given winter, it is essentially the peak demand which would be observed given that winter's underlying demand patterns and a typical winter's weather. This is distinct from the actual observed peak demand which depends on that particular winter's weather as well as underlying patterns. A formal definition may be found in the glossary of [8].

[3]The results are presented using data from (at the time of writing) two years ago, as the margin of conventional plant over peak demand was then deemed to be economically sustainable in the long run, making calculations representative of the long-run risk levels. At present, the plant margin is much higher due to gas generation being commissioned in anticipation of a large capacity of coal generation retiring around 2013.

imum capacity available. In combining the units, their availability states are assumed to be independent, and the availability probabilities for each class of generating unit are taken from [9]; in a small number of cases, the maximum contribution from each station is capped due to finite network capacity or emissions constraints. When the individual unit or station distributions are convolved explicitly, this is usually called the Capacity Outage Probability Table (COPT) method. In all examples, the distributions of available conventional capacity at different times are assumed to be identically distributed. This is reasonable if there is little planned maintenance at times when risk is high, and hence all units that are mechanically available are available to generate if required. This assumption of identical conventional plant distributions across relevant times is usually made in practical capacity value calculations.

The wind and demand data have the great benefit of being metered data, so do they do not include errors arising from conversion of meteorological records to these power system quantities. There are however a number of other factors which should be included in a quantitative projection of future risk levels in GB, such as how underlying demand patterns may change, how the distribution of available wind LF will change as the geographical distribution of wind capacities changes, the future conventional plant mix and its availability properties, and the uncertainty in all of these. The results presented should therefore *not* be treated as quantitative assessments of risk levels on the real GB system. Nevertheless, this data will prove very suitable in illustrating the application of the new theory to a range of risk calculation structures which are in widespread use.

The marginal distributions of the wind and demand data are visualised in Fig. 1 (in what is essentially cumulative form). The observed distribution of demands decays slightly faster than exponential at high demands; this plot also illustrates how (by definition) demand exceeds ACS peak in some winters. The wind data is visualised by plotting the mean load factor across all hours with demand above a given level. While there is much variability about this trend, this illustrates how the typical load factor deteriorates from typical winter values of about 30%, to low 20s % as demand increases from 90% to 95% of ACS peak. Above 95% of ACS peak, the quantity of data is too small to permit a robust statistical analysis; more detailed discussion of these limited data issues may be found in [10].

The probability distribution of available conventional capacity is plotted in Fig. 2. The left tail of this distribution decays more rapidly than the right tail of the distribution of demands, indicating that in GB the calculated LOLE will be dominated by times of extreme demand.

3.2 Fixed demand, risk index LOLP

In this first example, a fixed peak demand d is assumed. The theory of Section 2 applies directly to this case, with margin $M = X - d$. The risk index is then the LOLP at time of annual peak demand. An annual peak demand risk assessment approach has historically been used in Great Britain, for instance in the System Operator's generation adequacy assessment contained in its annual Winter Outlook [9], or in the pre-privatisation generation planning standard [11].

The data are used as described in Section 2.5. Figure 3 compares the direct solution to (1), and the marginal approximation (3), for a range of fixed peak demands. The probability distribution of available wind load factor used is the empirical distribution of historic load factors across hours where the demand was above 90% of ACS peak. The marginal approximation is very good at small installed wind capacities. These results illustrate the theoretical result that for small wind penetrations the percentage ELCC is the mean wind load factor of 29.2%, and for higher penetrations the percentage ELCC decreases as the increased variability of supply due to the additional generation becomes more important. This is consistent with observations made in [3], [4], and elsewhere.

The new theory also explains (in terms of differences in $f'_M(0)/f_M(0)$) the increase in calculated ELCC as the demand is increased This effect has also been observed in many other studies, for instance [4], and is usually explained intuitively in terms of additional generation being more valuable to the system when the risk level is higher. The new result (3) demonstrates that whether the ELCC increases or decreases as the demand is increased actually depends on the sign of

$$\frac{\mathrm{d}}{\mathrm{d}m} \frac{f'_M(m)}{f_M(m)}\bigg|_{M=0}.$$

The values of f'/f for the plots in Fig. 3 are given in Table 1, and the standard deviation of the distribution of the wind load factor is 23.4%. If the derivative of f'/f is negative (as it is in examples presented here), which is equivalent to the distribution being locally lighter-tailed than an exponential distribution, then the correction due to the variation of the additional generation decreases as the fixed demand is increased, and the ELCC increases. However, if the distribution of margin were locally heavier-tailed than an exponential distribution, the change in the ELCC would be in the opposite direction. The intuitive explanation in terms of 'higher risk, higher value of new generation' is thus not correct in general (though the circumstances under which it is not correct are apparently seldom realised in practice).

Another example of the effect of $f'_M(0)/f_M(0)$ is the difference between the ELCC calculation using a COPT-based distribution of available conventional capacity (i.e. explicit convolution of the distributions for individual generating units), and the ELCC using a Gaussian approximation, in Fig. 4. The reduction of capacity values below the mean is greater for the Gaussian approximation, for which $f'_M(0)/f_M(0) = 0.00124$ as compared to 0.00102 for the COPT calculation.

All of these examples illustrate how the calculated capacity value is a function of the pre-existing system background, and not just of the additional generation.

3.3 Risk index LOLE, demand and wind independent

In this example, the distribution of available wind LF is based on the empirical distribution of historic LFs over hours where demand was above 90% of ACS peak, and is assumed independent of demand. The historic demands are scaled to a common ACS peak of 60 GW, and the distribution of demand D is the empirical distribution of rescaled historic demands. In addition, as described in Section 3.1, the distributions of available conventional capacity at relevant times are assumed to be identical.

The distribution of $M = X - D$ is calculated by convolution of the distributions for D and available conventional capacity X. The 'snapshot' expression (1) may then be applied, with $\mathbf{P}(M + Y \leq 0)$ corresponding to LOLE divided by the number of time periods in the sum defining the LOLE.

Fig. 4 illustrates that once more (1) is a good approximation for small installed wind capacities. It may further be observed that in this variable demand case the ELCC results are quite similar to those using a fixed demand of 60 GW (for the former, $f'_M(0)/f'_M(0)$ is 0.00111, and for the latter 0.00103.)

3.4 Risk index LOLE, demand and wind non-independent

The results of a hindcast-based ELCC calculation — as discussed in Section 2.5 and requiring no prior assumptions about the distributions of D and W or their dependence structure — are shown in Fig. 4. In comparison with the results based on a wind distribution independent of demand, and based on all hours above 90% of ACS peak, the hindcast results are somewhat lower. This is because, as discussed in Section 3.1, the estimated LOLE is dominated by demands near ACS peak, at which the typical wind load factor is lower than across all times when demand is above 90% of ACS peak.

The results from previous sections would more closely match those from this hindcast result if the cutoff were a higher percentage of ACS peak, for instance 95%. However, due to the paucity of data available (and other issues described in Section 3.1), the authors make no claim that the results presented accurately represent true risk levels or wind's true contribution in Great Britain; the examples presented are instead intended to illustrate the results of Section 2[4].

The relevant theory is developed in Sections 2.4 and 2.5. The leading order expression for the hindcast ELCC is then

$$\nu_Y^{\mathrm{ELCC}} = \frac{\sum_t y_t f_X(d_t)}{\sum_t f_X(d_t)} \tag{18}$$

and Fig. 4 includes a comparison with this result. The next order in the marginal ELCC result is not presented for the hindcast calculation, as it requires explicit estimation of $\sigma_{Y|D}$.

[4]Furthermore, the choice of 90% cutoff emphasises this point that the results should not be interpreted as a quantitative assessment of generation adequacy risk levels in GB. It is possible that a more practically realistic cutoff such as 95% might serve only to create confusion on this point.

Section 2.3 demonstrates that the marginal capacity value result in the non-independent case is dominated by demand levels at which the product $f_X(x)f_D(x)$ is high. In GB, it is extremely unlikely that $M = X - D$ will be below zero unless the demand is extremely high, as the distribution of X is narrow compared to the scale on which that for D varies; this corresponds to $f_X(x)f_D(x)$ being large only when X is near peak demand. In a smaller system, however, where there are fewer conventional units, the distribution of available conventional capacity will typically be broader relative to peak demand; as a consequence, it would not be necessary to have a truly extreme demand in order for M to be negative, and hence the wind resource at lower demands will be significant in determining the ELCC. This provides a further example of the calculated capacity value of the additional wind generation depending on the pre-existing background of demand and conventional plant.

4 Conclusions

This paper presents the probability theory of the capacity value of additional generation in power systems. Without any assumptions regarding distribution shapes, a closed-form expression for the effective load carrying capability or equivalent firm capacity of additional generation may be derived. The leading order term in this expression is simply the mean of the distribution for available additional capacity, and the next order term is a correction proportional to the variance of the distribution for available conventional capacity.

The theory may be used to explain and confirm the generality of various well-known properties of capacity value results. Of particular note is the common observation that if the distribution for demand is shifted so as to increase the calculated risk, then the capacity value of additional generation increases. This is usually explained in terms of additional generation being more valuable to the system when the risk level is higher. However, the new theory demonstrates that this effect will only be observed when the distribution of the margin of available existing capacity over demand is locally light-tailed near zero margin; if the distribution were locally heavy-tailed, then the calculated capacity value would decrease as risk is increased in this manner.

The theory extends in a natural way to cases where the available additional capacity and demand are not independent (as may well be the case when the additional generation represents wind capacity). In this case, the leading order result may be verified using a standard hindcast calculation.

Acknowledgements

The authors are grateful to National Grid for sharing metered wind output data, and for many valuable discussions including their hosting Dr. Dent's secondment to the Control Centre. They acknowledge further important discussions with colleagues at their own universities and the Electricity Research Centre at University College Dublin.

Funding

Dr. Zachary's work was supported by a Scottish Funding Council SPIRIT award: Mathematical Methods to Support the Integration of Renewables into the Electricity Network. Dr. Dent's work was supported by the Supergen FlexNet consortium as part of the UK Research Councils Energy Programme under grant [EP/E04011X/1], and by an associated Engineering and Physical Sciences Research Council Knowledge Transfer Secondment to National Grid under grant [EP/H500340/1].

Declaration of conflicting interests

The authors declare that there are no conflicts of interest.

A Nomenclature

Throughout, capital letters denote random variables.

X Available existing capacity.

Y Available additional capacity.

D Demand.

M Margin of existing capacity over demand.

ν_Y^{ELCC} Effective Load Carrying Capability (ELCC) of additional generation Y.

ν_Y^{EFC} Equivalent Firm Capacity (EFC) of additional generation Y.

ACS Average Cold Spell. ACS peak is the standard measure of underlying demand level in Great Britain.

COPT Capacity Outage Probability Table. Usual name given to derivation of distribution of available conventional capacity by convolving distributions for individual units or stations.

LF Load factor, a generator's output as a proportion of maximum.

LOLP Loss of Load Probability, the probability of insufficient generating capacity being available to meet demand at any instant in time.

LOLE Loss of Load Expectation, the expected number of periods in a time window in which capacity is insufficient available capacity to meet demand, or equivalently the sum over periods of LOLP.

References

[1] Doherty R. All Island Grid Study Workstream 2A: High Level Assessment of Suitable Generation Portfolios for the All-Island System in 2020; 2008. Available at http://www.dcenr.gov.ie/Energy/North-South+Co-operation+in+the+Energy+Sector/All+Island+Electricity+Grid+Study.htm.

[2] National Renewable Energy Laboratory. Eastern Wind Integration and Transmission Study; 2010. Available at http://www.nrel.gov/wind/systemsintegration/ewits.html.

[3] Keane A, Milligan M, Dent CJ, Hasche B, C DAnnunzio, Dragoon K, et al.. Capacity Value of Wind Power;. Report of the IEEE PES Task Force on Capacity Value of Wind Power, accepted for IEEE Trans. Power Syst.

[4] Amelin M. Comparison of Capacity Credit Calculation Methods for Conventional Power Plants and Wind Power. IEEE Trans Power Syst. 2009 May;24(2):685–691.

[5] Dragoon K, Dvortsov V. Z-Method for Power System Resource Adequacy Applications. IEEE Trans Power Syst. 2006 May;21(2).

[6] National Grid Company, personal communication;.

[7] National Grid: Demand Data;. Available at http://www.nationalgrid.com/uk/Electricity/Data/Demand+Data/.

[8] Great Britain National Electricity Transmission System Security and Quality of Supply Standard, Issue 2.0; 24 June 2009. Available from http://www.nationalgrid.com/uk/Electricity/Codes/gbsqsscode/DocLibrary/.

[9] National Grid. Winter Outlook Report 2010/11; 7 October 2010. Available from http://www.nationalgrid.com/uk/Gas/TYS/outlook/.

[10] Zachary S, Dent CJ, Brayshaw DJ. Challenges in Quantifying Wind Generation's Contribution to Securing Peak Demand. In: IEEE Power and Energy Society General Meeting; 2011. .

[11] Dent CJ, Bell KRW, Richards AW, Zachary S, Eager D, Harrison GP, et al. The Role of Risk Modelling in the Great Britain Transmission Planning and Operational Standards. In: International Conference on Probabilistic Methods Applied to Power Systems; 2010. .

Figure 1: Visualisation of the historic wind and demand data used in the Great Britain examples.

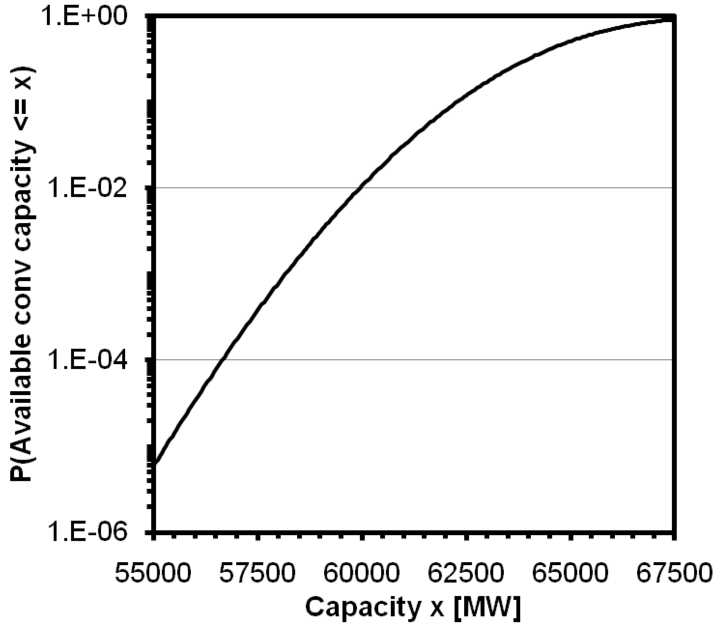

Figure 2: Probability distribution of available conventional capacity.

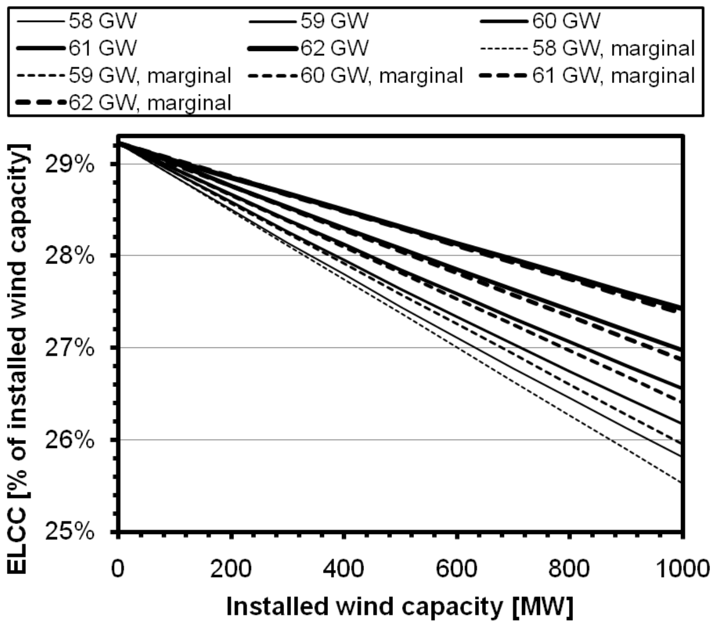

Figure 3: Wind ELCC in the GB test system for a range of fixed peak demands. Solid lines indicate a direct solution to (1), and dashed lines the marginal approximation (3).

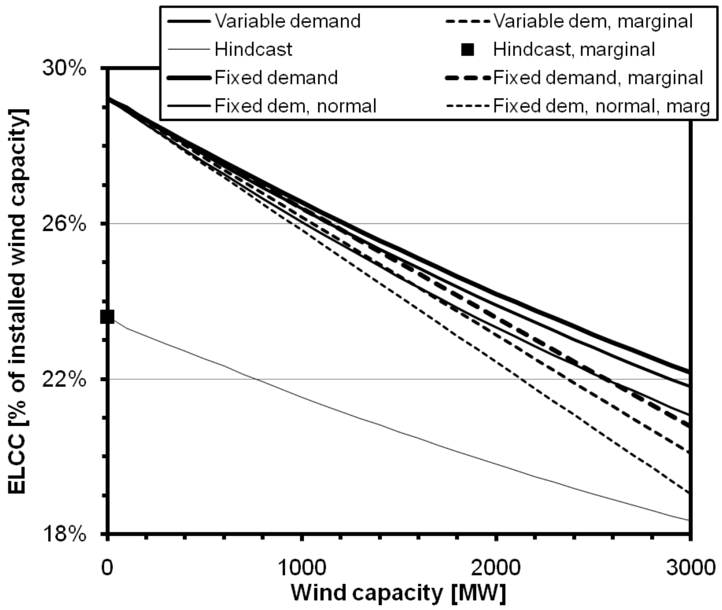

Figure 4: Wind ELCC in the GB test system for fixed peak demand of 60 GW (with both COPT-based and Gaussian approximation for conventional plant), an LOLE calculation with wind and demand independent, and a hindcast calculation. Again, solid lines indicate a direct solution to (1), and dashed lines the marginal approximation (3).

ACS peak [GW]	58	59	60	61	62
(f'/f) $[\times 10^3]$	1.35	1.19	1.023	0.86	0.67

Table 1: $f'_M(0)/f_M(0)$ for the peak demand levels shown in Fig. 3.